(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 October 2002 (31.10.2002)

PCT

(10) International Publication Number
WO 02/086867 A1

(51) International Patent Classification[7]: G10L 19/02 // 21/02

(21) International Application Number: PCT/SE02/00485

(22) International Filing Date: 14 March 2002 (14.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0101408-3 23 April 2001 (23.04.2001) SE

(71) Applicant: TELEFONAKTIEBOLAGET L M ERICSSON (publ) [SE/SE]; S-126 25 Stockholm (SE).

(72) Inventors: NILSSON, Mattias; Hjortronvägan 109, 1 tr, S-196 35 Kungsängen (SE). KLEIJN, Bastiaan; Tallåsvägen 11, S-182 75 Stocksund (SE).

(74) Agents: BERGLUND, Stefan et al.; Bjerkéns Patentbyrå KB, Östermalmsgatan 58, S-114 50 Stockhlom (SE).

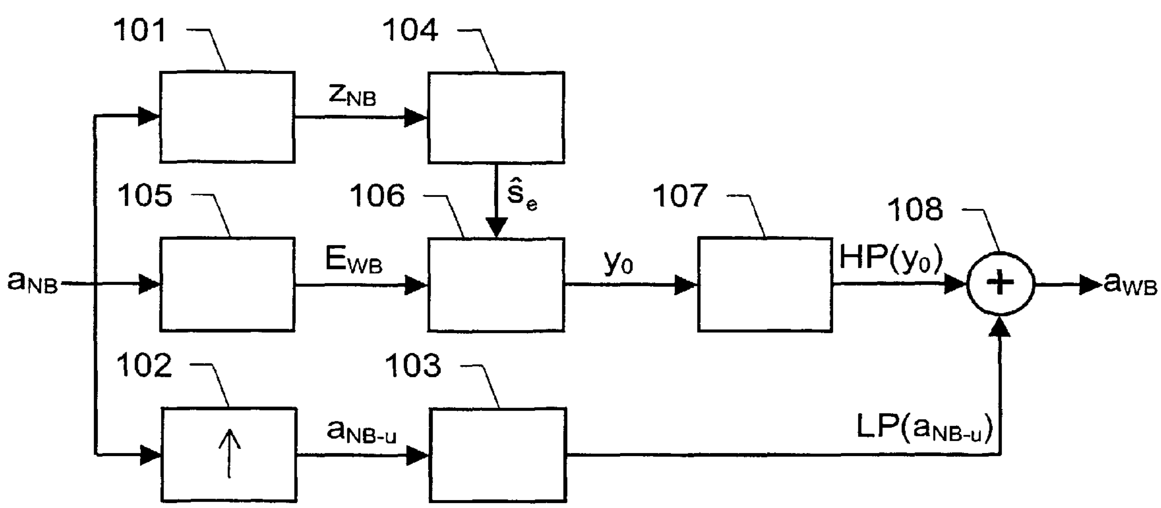
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: BANDWIDTH EXTENSION OF ACOUSIC SIGNALS



(57) Abstract: The present invention relates to a solution for improving the perceived sound quality of a decoded acoustic signal. THe improvement is accomplised by means of extending the spectrum of a received narrow-band acoustic signal (a_{NB}). According to the invention , a wide-band acoustic signal (a_{WB}) is produced by extracting at least one essential attribute (Z_{NB}) from the narrow-band acoustic signal (a_{NB}). Parameters, e.g. representing signal energies, with respect to wide-band frequency components outside the spectrum (A_{NB}) of the narrow-band acoustic signal (a_{NB}) are estimated based on the at least one essential attribute (Z_{NB}).This estimation involves allocating a parameter value to a wide-band frequency component, based on a corresponding confidence level. For instance, a relatively high parameter value is allowed to be allocated to a frequency component if it has a comparatively high degree certainty. In contrast, a relatively low parameter value is only allowed to be allocated to a frequency component if it is associated with a comparatively low degree certainty.

WO 02/086867 A1

Bandwidth Extension of Acoustic Signals

THE BACKGROUND OF THE INVENTION AND PRIOR ART

The present invention relates generally to the improvement of the perceived sound quality of decoded acoustic signals. More particularly the invention relates to a method of producing a wide-band acoustic signal on basis of a narrow-band acoustic signal according to the preamble of claim 1 and a signal decoder according to the preamble of claim 24. The invention also relates to a computer program according to claim 22 and a computer readable medium according to claim 23.

Today's public switched telephony networks (PSTNs) generally low-pass filter any speech or other acoustic signal that they transport. The low-pass (or, in fact, band-pass) filtering characteristic is caused by the networks' limited channel band-width, which typically has a range from 0,3 kHz to 3,4 kHz. Such band-pass filtered acoustic signal is normally perceived by a human listener to have a relatively poor sound quality. For instance, a reconstructed voice signal is often reported to sound muffled and/or remote from the listener.

The trend in fixed and mobile telephony as well as in video-conferencing is, however, towards an improved quality of the acoustic source signal that is reconstructed at the receiver end. This trend reflects the customer expectation that said systems provide a sound quality, which is much closer to the acoustic source signal than what today's PSTNs can offer.

One way to meet this expectation is, of course, to broaden the frequency band for the acoustic source signal and thus convey more of the information being contained in the source signal to the receiver. For instance, if a 0 – 8 kHz acoustic signal (sampled at 16 kHz) were transmitted to the receiver, the naturalness of a human voice signal, which is otherwise lost in a standard phone call, would indeed be better preserved. However, increasing the bandwidth for each channel by more than a factor two would either reduce the transmission capacity to less than half or imply enormous costs for the network operators in order to expand the transmission resources by a corresponding factor. Hence, this solution is not attractive from a commercial point-of-view.

Instead, recovering at the receiver end, wide-band frequency components outside the bandwidth of a regular PSTN-channel based on the narrow-band signal that has passed through the PSTN constitutes a much more appealing alternative. The recovered wide-band frequency components may both lie in a low-band below the narrow-band (e.g. in a range 0,1 – 0,3 kHz) and in a high-band above the narrow-band (e.g. in a range 3,4 – 8,0 kHz).

Although the majority of the energy in a speech signal is spectrally located between 0 kHz and 4 kHz, a substantial amount of the energy is also distributed in the frequency band from 4 kHz to 8 kHz. The frequency resolution of the human hearing decreases rapidly with increasing frequencies. The frequency components between 4 kHz and 8kHz therefore require comparatively small amounts of data to model with a sufficient accuracy.

It is possible to extend the bandwidth of the narrow-band acoustic signal with a perceptually satisfying result, since the signal is presumed to be generated by a physical source, for instance, a human speaker. Thus, given a particular shape of the narrow-band, there are constraints on the signal properties

with respect to the wide-band shape. I.e. only certain combinations of narrow-band shapes and wide-band shapes are conceivable.

However, modelling a wide-band signal from a particular narrow-band signal is still far from trivial. The existing methods for extending the bandwidth of the acoustic signal with a high-band above the current narrow-band spectrum basically include two different components, namely: estimation of the high-band spectral envelope from information pertaining to the narrow-band, and recovery of an excitation for the high-band from a narrow-band excitation.

All the known methods, in one way or another, model dependencies between the high-band envelope and various features describing the narrow-band signal. For instance, a Gaussian mixture model (GMM), a hidden Markov model (HMM) or vector quantisation (VQ) may be utilised for accomplishing this modelling. A minimum mean square error (MMSE) estimate is then obtained from the chosen model of dependencies for the high-band spectral envelope provided the features that have been derived from the narrow-band signal. Typically, the features include a spectral envelope, a spectral temporal variation and a degree of voicing.

The narrow-band excitation is used for recovering a corresponding high-band excitation. This can be carried out by simply up-sampling the narrow-band excitation, without any following low-pass filtering. This, in turn, creates a spectral-folded version of the narrow-band excitation around the upper bandwidth limit for the original excitation. Alternatively, the recovery of the high-band excitation may involve techniques that are otherwise used in speech coding, such as multi-band excitation (MBE). The latter makes use of the fundamental frequency and the degree of voicing when modelling an excitation.

Irrespective of how the high-band excitation is derived, the estimated high-band spectral envelope is used for obtaining a desired shape of the recovered high-band excitation. The result thereof in turn forms a basis for an estimate of the high-band acoustic signal. This signal is subsequently high-pass filtered and added to an up-sampled and low-pass filtered version of the narrow-band acoustic signal to form a wide-band acoustic signal estimate.

Normally, the bandwidth extension scheme operates on a 20-ms frame-by-frame basis, with a certain degree of overlap between adjacent frames. The overlap is intended to reduce any undesired transition effects between consecutive frames.

Unfortunately, the above-described methods all have one undesired characteristic in common, namely that they introduce artefacts in the extended wide-band acoustic signals. Furthermore, it is not unusual that these artefacts are so annoying and deteriorate the perceived sound quality to such extent that a human listener generally prefers the original narrow-band acoustic signal to the thus extended wide-band acoustic signal.

SUMMARY OF THE INVENTION

The object of the present invention is therefore to provide an improved bandwidth extension solution for a narrow-band acoustic signal, which alleviates the problem above and thus produces a wide-band acoustic signal that has a significantly enhanced perceived sound quality. The above-indicated problem being associated with the known solutions is generally deemed to be due to an over-estimation of the wide-band energy (predominantly in the high-band).

According to one aspect of the invention the object is achieved by a method of producing a wide-band acoustic signal on basis of a narrow-band acoustic signal as initially described, which is characterised by allocating a parameter with respect to a

particular wide-band frequency component based on a corresponding confidence level.

According to a preferred embodiment of the invention, a relatively high parameter value is thereby allowed to be allocated to a frequency component if the confidence level indicates a comparatively high degree certainty. In contrast, a relatively low parameter value is allowed to be allocated to a frequency component if the confidence level indicates a comparatively low degree certainty.

According to one embodiment of the invention, the parameter directly represents a signal energy for one or more wide-band frequency components. However, according to an alternative embodiment of the invention, the parameter only indirectly reflects a signal energy. The parameter then namely represents an upper-most bandwidth limit of the wide-band acoustic signal, such that a high parameter value corresponds to a wide-band acoustic signal having a relatively large bandwidth, whereas a low parameter value corresponds to a more narrow bandwidth of the wide-band acoustic signal.

According to a further aspect of the invention the object is achieved by a computer program directly loadable into the internal memory of a computer, comprising software for performing the method described in the above paragraph when said program is run on a computer.

According to another aspect of the invention the object is achieved by a computer readable medium, having a program recorded thereon, where the program is to make a computer perform the method described in the penultimate paragraph above.

According to still another aspect of the invention the object is achieved by a signal decoder for producing a wide-band acoustic signal from a narrow-band acoustic signal as initially described, which is characterised in that the signal decoder is

arranged to allocate a parameter to a particular wide-band frequency component based on a corresponding confidence level.

According to a preferred embodiment of the invention, the decoder thereby allows a relatively high parameter value to be allocated to a frequency component if the confidence level indicates a comparatively high degree certainty, whereas it allows a relatively low parameter value to be allocated to a frequency component whose confidence level indicates a comparatively low degree certainty.

In comparison to the previously known solutions, the proposed solution significantly reduces the amount of artefacts being introduced when extending a narrow-band acoustic signal to a wide-band representation. Consequently, a human listener perceives a drastically improved sound quality. This is an especially desired result, since the perceived sound quality is deemed to be a key factor in the success of future telecommunication applications.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is now to be explained more closely by means of preferred embodiments, which are disclosed as examples, and with reference to the attached drawings.

Figure 1 shows a block diagram over a general signal decoder according to the invention,

Figure 2 exemplifies a spectrum of a typical acoustic source signal in the form of a speech signal,

Figure 3 exemplifies a spectrum of the acoustic source signal in figure 2 after having been passed through a narrow-band channel,

Figure 4 exemplifies a spectrum of the acoustic signal corresponding to the spectrum in figure 3 after having been extended to a wide-band acoustic signal according to the invention,

Figure 5 shows a block diagram over a signal decoder according to an embodiment of the invention,

Figure 6 illustrates a narrow-band frame format according to an embodiment of the invention,

Figure 7 shows a block diagram over a part of a feature extraction unit according to an embodiment of the invention,

Figure 8 shows a graph over an asymmetric cost-function, which penalizes over-estimates of an energy-ratio between the high-band and the narrow-band according to an embodiment of the invention, and

Figure 9 illustrates, by means of a flow diagram, a general method according to the invention.

DESCRIPTION OF PREFERRED EMBODIMENTS OF THE INVENTION

Figure 1 shows a block diagram over a general signal decoder according to the invention, which aims at producing a wide-band acoustic signal a_{WB} on basis of a received narrow-band signal a_{NB}, such that the wide-band acoustic signal a_{WB} perceptually resembles an estimated acoustic source signal a_{source} as much as possible. It is here presumed that the acoustic source signal a_{source} has a spectrum A_{source}, which is at least as wide as the bandwidth W_{WB} of the wide-band acoustic signal a_{WB} and that the wide-band acoustic signal a_{WB} has a wider spectrum A_{WB} than the spectrum A_{NB} of the narrow-band acoustic signal a_{NB}, which has been transported via a narrow-band channel that has a

bandwidth W_{NB}. These relationships are illustrated in the figures 2–4. Moreover, the bandwidth W_{WB} may be sub-divided into a low-band W_{LB} including frequency components between a low-most bandwidth limit f_{Wl} below a lower bandwidth limit f_{Nl} of the narrow-band channel and the lower bandwidth limit f_{Nl} respective a high-band W_{HB} including frequency components between an upper-most bandwidth limit f_{Wu} above an upper bandwidth limit f_{Nu} of the narrow-band channel and the upper bandwidth limit f_{Nu}.

The proposed signal decoder includes a feature extraction unit 101, an excitation extension unit 105, an up-sampler 102, a wide-band envelope estimator 104, a wide-band filter 106, a low-pass filter 103, a high-pass filter 107 and an adder 108. The feature extraction unit's 101 function will be described in the following paragraph, however, the remaining units 102 – 108 will instead be described with reference to the embodiment of the invention shown in figure 5.

The signal decoder receives a narrow-band acoustic signal a_{NB}, either via a communication link (e.g. in PSTN) or from a storage medium (e.g. a digital memory). The narrow-band acoustic signal a_{NB} is fed in parallel to the feature extraction unit 101, the excitation extension unit 105 and the up-sampler 102. The feature extraction unit 101 generates at least one essential feature z_{NB} from the narrow-band acoustic signal a_{NB}. The at least one essential feature z_{NB} is used by the following wide-band envelope estimator 104 to produce a wide-band envelope estimation $\hat{s}_e$. A Gaussian mixture model (GMM) may, for instance, be utilised to model the dependencies between the narrow-band feature vector z_{NB} and a wide-/high-band feature vector z_{WB}. The wide-/high band feature vector z_{WB} contains, for instance, a description of the spectral envelope and the logarithmic energy-ratio between the narrow-band and a wide-/high-band. The narrow-band feature vector z_{NB} and the wide-/high-band feature vector z_{WB} are combined into a joint feature vector $z=[z_{NB}, z_{WB}]$. The GMM models a joint probability density function $f_Z(z)$ of a random variable feature vector Z, which can

be expressed as:

$$f_Z(z)=\sum_{m=1}^{M}\alpha_m f_Z(z\mid\theta_m)$$

where M represents a total number of mixture components, α_m is a weight factor for a mixture number m and $f_Z(z|\theta_m)$ is a multivariate Gaussian distribution, which in turn is described by:

$$f_Z(z\mid\theta_m)=\frac{1}{(2\pi)^{d/2}\left|C_m\right|^{1/2}}\exp\left(-\frac{1}{2}(z-\mu_{zm})^{\tau}C_m^{-1}(z-\mu_{zm})\right)$$

where μ_m represents a mean vector and C_m is a covariance matrix being collected in the variable $\theta_m=\{\mu_m, C_m\}$ and d represents a feature dimension. According to an embodiment of the invention the feature vector z has 22 dimensions and consists of the following components:

a narrow-band spectral envelope, for instance modelled by 15 linear frequency cepstral coefficients (LFCCs), i.e. $x=\{x_1, \ldots, x_{15}\}$,

a high-band spectral envelope, for instance modelled by 5 linear frequency cepstral coefficients, i.e. $y=\{y_1, \ldots, y_5\}$,

an energy-ratio variable g denoting a difference in logarithmic energy between the high-band and the narrow-band, i.e. $g=y_0-x_0$, where y_0 is the logarithmic high-band energy and x_0 is the logarithmic narrow-band energy, and

a measure representing a degree of voicing r. The degree of voicing r may, for instance, be determined by localising a maximum of a normalised autocorrelation function within a lag range corresponding to 50 – 400 Hz.

According to an embodiment of the invention, the weight factor α_m and the variable θ_m for m=1, ..., M are obtained by applying the so-called estimate-maximise (EM) algorithm on a training set being extracted from the so-called TIMIT-database (TIMIT = Texas Instruments / Massachusetts Institute of Technology).

The size of the training set is preferably 100 000 non-overlapping 20 ms wide-band signal segments. The features z are then extracted from the training set and their dependencies are modelled by, for instance, a GMM with 32 mixture components (i.e. M=32).

Figure 5 shows a block diagram over a signal decoder according to an embodiment of the invention. By way of introduction, the over all working principle of the decoder is described. Next, the operation of the specific units included in the decoder will be described in further detail.

The signal decoder receives a narrow-band acoustic signal a_{NB} in the form of segments, which each has a particular extension in time T_f, e.g. 20 ms. Figure 6 illustrates an example narrow-band frame format according to an embodiment of the invention, where a received narrow-band frame n is followed by subsequent frames n+1 and n+2. Preferably, adjacent segments overlap each other to a specific extent T_o, e.g. corresponding to 10 ms. According to an embodiment of the invention, 15 cepstral coefficients x and a degree of voicing r are repeatedly derived from each incoming narrow-band segment n, n+1, n+2 etc.

Then, an estimate of an energy-ratio between the narrow-band and a corresponding high-band is derived by a combined usage of an asymmetric cost-function and an a-posteriori distribution of energy-ratio based on the narrow-band shape (being modelled by the cepstral coefficients x) and the narrow-band voicing parameter (described by the degree of voicing r). The asymmetric cost-function penalizes over-estimates of the energy-ratio more than under-estimates of the energy-ratio. Moreover, a narrow a-posteriori distribution results in less penalty on the energy-ratio than a broad a-posteriori distribution. The energy-ratio estimate, the narrow-band shape x and the degree of voicing r together form a new a-posteriori distribution of the high-band shape. An MMSE estimate of the high-band envelope is also computed on basis of the energy-

ratio estimate, the narrow-band shape x and the degree of voicing r. Subsequently, the decoder generates a modified spectral-folded excitation signal for the high-band. This excitation is then filtered with the energy-ratio controlled high-band envelope and added to the narrow-band to form a wide-band signal a_{WB}, which is fed out from the decoder.

The feature extraction unit 101 receives the narrow-band acoustic signal a_{NB} and produces in response thereto at least one essential feature $z_{NB}(r, c)$ that describes particular properties of the received narrow-band acoustic signal a_{NB}. The degree of voicing r, which represents one such essential feature $z_{NB}(r, c)$, is determined by localising a maximum of a normalised autocorrelation function within a lag range corresponding to 50 – 400 Hz. This means that the degree of voicing r may be expressed as:

$$r = \max_{20 \le r \le 160} \frac{\sum_{n=0}^{N-1} s(n)s(n+\tau)}{\sqrt{\sum_{k=0}^{N-1} s(k)^2 \sum_{i=0}^{N-1} s(i+\tau)^2}}$$

where s=s(1), ..., s(160) is a narrow-band acoustic segment having a duration of T_f (e.g. 20ms) being sampled at, for instance, 8 kHz.

The spectral envelope c is here represented by LFCCs. Figure 7 shows a block diagram over a part of the feature extraction unit 101, which is utilised for determining the spectral envelope c according to this embodiment of the invention.

A segmenting unit 101a separates a segment s of the narrow-band acoustic signal a_{NB} that has a duration of T_f = 20 ms. A following windowing unit 101b windows the segment s with a window-function w, which may be a Hamming-window. Then, a transform unit 101c computes a corresponding spectrum S_W by means of a fast Fourier transform, i.e. S_W = FFT(w·s). The envelope S_E of the spectrum S_W of the windowed narrow-band acoustic signal a_{NB} is obtained by convolving the spectrum S_W

with a triangular window W_T in the frequency domain, which e.g. has a bandwidth of 100 Hz, in a following convolution unit 101d. Thus, $S_E = S_W * W_T$.

A logarithm unit 101e receives the envelope S_E and computes a corresponding logarithmic value $S_E^{\log}$ according to the expression:

$$S_E^{\log} = 20\log_{10}(S_E)$$

Finally, an inverse transform unit 101f receives the logarithmic value $S_E^{\log}$ and computes an inverse fast Fourier transform thereof to represent the LFCCs, i.e.:

$$c = \mathrm{IFFT}(S_E^{\log})$$

where c is a vector of linear frequency cepstral coefficients. A first component c_0 of the vector c constitutes the log energy of the narrow-band acoustic segment s. This component c_0 is further used by a high-band shape reconstruction unit 106a and an energy-ratio estimator 104a that will be described below. The other components $c_1, \ldots, c_{15}$ in the vector c are used to describe the spectral envelope x, i.e. $x = [c_1, \ldots, c_{15}]$.

The energy-ratio estimator 104a, which is included in the wide-band envelope estimator 104, receives the first component c_0 in the vector of linear frequency cepstral coefficients c and produces, on basis thereof, plus on basis of the narrow-band shape x and the degree of voicing r an estimated energy-ratio $\hat{g}$ between the high-band and the narrow-band. In order to accomplish this, the energy-ratio estimator 104a uses a quadratic cost-function, as is common practice for parameter estimation from a conditioned probability function. A standard MMSE estimate $\hat{g}_{MMSE}$ is derived by using the a-posteriori distribution of the energy-ratio given the narrow-band shape x and the degree of voicing r together with the quadratic cost-function, i.e.:

$$
\begin{aligned}
\hat{g}_{MMSE} &= \arg\min_{\hat{y}} \int_{\Omega_g} (\hat{g}-g)^2 f_{G|XR}(g \mid x,r)dg \\
&= E[G|X=x,\ R=r] \\
&= \int_{\Omega_g} g \frac{\sum_{m=1}^{M} \alpha_m f_{GXR}(g,x,r \mid \theta_m)}{\sum_{k=1}^{M} \alpha_k f_{XR}(x,r \mid \theta_k)} dg \\
&= \sum_{m=1}^{M} \frac{\alpha_m f_{XR}(x,r \mid \theta_m)}{\sum_{k=1}^{M} \alpha_k f_{XR}(x,r \mid \theta_k)} \int_{\Omega_g} g f_{G|XR}(g \mid x,r,\theta_m)dg \\
&= \sum_{m=1}^{M} w_m(x,r) \int_{\Omega_g} g f_{G|XR}(g \mid x,r,\theta_m)dg \\
&= \sum_{m=1}^{M} w_m(x,r) \int_{\Omega_g} g f_{G}(g \mid \theta_m)dg \\
&= \sum_{m=1}^{M} w_m(x,r)\mu_{y_m}
\end{aligned}
$$

where in the second last step, the fact is used, that each individual mixture component has a diagonal covariance matrix and, thus, independent components. Since an over-estimation of the energy-ratio is deemed to result in a sound that is perceived as annoying by a human listener, an asymmetric cost-function is used instead of a symmetric ditto. Such function is namely capable of penalising over-estimates more that under-estimates of the energy-ratio. Figure 8 shows a graph over an exemplary asymmetric cost-function, which thus penalizes over-estimates of the energy-ratio. The asymmetric cost-function in figure 8 may also be expressed as:

$$C = bU(\hat{g}-g) + (\hat{g}-g)^2$$

where bU(•) represents a step function with an amplitude b. The amplitude b can be regarded as a tuning parameter, which provides a possibility to control the degree of penalty for the over-estimates. The estimated energy-ratio $\hat{g}$ can be expressed as:

$$\hat{g} = \arg\min_{g} \int_{\Omega_g} (bU(\hat{g} - g) + (\hat{g} - g)^2) f_{G|XR}(g \mid x, r) dg$$

The estimated energy-ratio $\hat{g}$ is found by differentiating the right-hand side of the expression above and set it equal to zero. Assuming that the order of differentiation and integration may be interchanged the derivative of the above expression can be written as:

$$\sum_{m=1}^{M} w_m(x,r) \int_{\Omega_g} (b\delta(\hat{g} - g) + 2(\hat{g} - g)) f_G(g \mid \theta_m) dg = 0,$$

$$\sum_{m=1}^{M} w_m(x,r) b f_G(\hat{g} \mid \theta_m) + 2\hat{g} - 2\sum_{m=1}^{M} w_m(x,r) \mu_{y_m} = 0,$$

which in turn yields an estimated energy-ratio $\hat{g}$ as:

$$\hat{g} = \sum_{m=1}^{M} w_m(x,r) \mu_{y_m} - \frac{b}{2} \sum_{m=1}^{M} w_m(x,r) f_G(\hat{g} \mid \theta_m)$$

The above equation is preferably solved by a numerical method, for instance, by means of a grid search. As is apparent from the above, the estimated energy-ratio $\hat{g}$ depends on the shape posterior distribution. Consequently, the penalty on the MMSE estimate $\hat{g}_{MMSE}$ of the energy-ratio depends on the width of the posterior distribution. If the a-posteriori distribution $f_{G|XR}(g|x,r)$ is narrow, this means that the MMSE estimate $\hat{g}_{MMSE}$ is more reliable than if the a-posteriori distribution is broad. The width of the a-posteriori distribution can thus be seen as a confidence level indicator.

Other parameters than LFCCs can be used as alternative representations of the narrow-band spectral envelope x. Line Spectral Frequencies (LSF), Mel Frequency Spectral Coefficients (MFCC), and Linear Prediction Coefficients (LPC) constitute such alternatives. Furthermore, spectral temporal variations can be incorporated into the model either by including spectral derivatives in the narrow-band feature vector z_{NB} and/or by changing the GMM to a hidden Markov model (HMM).

Moreover, a classification approach may instead be used to express the confidence level. This means that a classification error is exploited to indicate a degree of certainty for a high-band estimate (e.g. with respect to energy y_0 or shape x).

According to an embodiment of the invention, it is presumed that the underlying model is GMM. A so-called Bayes classifier can then be constructed to classify the narrow-band feature vector z_{NB} into one of the mixture components of the GMM. The probability that this classification is correct can also be computed. Said classification is based on the assumption that the observed narrow-band feature vector z was generated from only one of the mixture components in the GMM. A simple scenario of a GMM that models the distribution of a narrow-band feature z using two different mixture components s_1; s_2 (or states) is shown below.

$$f_Z(z)=f_{Z,S}(z,s_1) + f_{Z,S}(z,s_2)$$

Suppose a vector z_0 is observed and the classification finds that the vector most likely originates from a realisation of the distribution in state s_1. Using Bayes rule, the probability $P(S=s_1|Z=z_0)$ that the classification was correct can be computed as:

$$P(S=s_1|Z=z_0) = \lim_{\Delta\to 0} P\left(S = s_1 \mid z_0 - \frac{\Delta}{2} < Z < z_0 + \frac{\Delta}{2}\right)$$

$$= \lim_{\Delta\to 0} \frac{\int_{z_0-\frac{\Delta}{2}}^{z_0+\frac{\Delta}{2}} f_{Z|S}(z \mid s_1)dz \cdot P(s_1)dz}{\int_{z_0-\frac{\Delta}{2}}^{z_0+\frac{\Delta}{2}} f_{Z|S}(z \mid s_1) \cdot P(s_1) + f_{Z|S}(z \mid s_2) \cdot P(s_2)dz}$$

$$= \frac{f_{Z|S}(z_0 \mid s_1) \cdot P(s_1)}{f_{Z|S}(z_0 \mid s_1) \cdot P(s_1) + f_{Z|S}(z_0 \mid s2) \cdot P(s_2)}$$

The probability of a correct classification can then be regarded as a confidence level. It can thus also be used to control the

energy (or shape) of the bandwidth extended regions W_{LB} and W_{HB} of the wide-band acoustic signal a_{WB}, such that a relatively high energy is allocated to frequency components being associated with a confidence level that represents a comparatively high degree certainty, and a relatively low energy is allocated to frequency components if the confidence level being associated with a confidence level that represents a comparatively low degree certainty.

The GMM is typically trained by means of an estimate-maximise (EM) algorithm in order to find the maximum likelihood estimate of the unknown, however, fixed parameters of the GMM given the observed data. According to an alternative embodiment of the invention, the unknown parameters of the GMM are instead themselves regarded as stochastic variables. A model uncertainty may also be incorporated by including a distribution of the parameters into the standard GMM. Consequently, the GMM would be a model of the joint distribution $f_{Z,\Theta}(z,\theta)$ of feature vectors z and the underlying parameters θ, i.e.:

$$f_{Z,\Theta}(z,\theta) = \sum_{m=1}^{M} \alpha_m f_{Z|\Theta}(z \mid \theta) f_{\Theta}(\theta)$$

The distribution $f_{Z,\Theta}(z,\theta)$ is then used to compute the estimates of the high-band parameters. For instance, as will be shown in further detail below, the expression for calculating the estimated energy-ratio $\hat{g}$, when using a proposed asymmetric cost-function, is:

$$\hat{g} = \arg\min_{g} \int_{\Omega_g} (bU(\hat{g} - g) + (\hat{g} - g)^2) f_{G|XR}(g \mid x, r) dg$$

An incorporation of the model uncertainty for the estimated energy-ratio $\hat{g}$ results in the expression:

$$\hat{g} = \arg\min_{g} \int_{\Omega_\theta} \int_{\Omega_g} (bU(\hat{g} - g) + (\hat{g} - g)^2) f_{G|XR}(g \mid x, r, \theta) f_{\Theta}(\theta) dg d\theta$$

Whenever the distribution $f_\Theta(\theta)$ and/or the distribution $f_{G|XR}(g|x,r,\theta)$ are broad, this will be interpreted as an indicator of a comparatively low confidence level, which in turn will result in a relatively low energy being allocated to the corresponding frequency components. Otherwise, (*i.e.* if both distributions $f_\Theta(\theta)$ and $f_{G|XR}(g|x,r,\theta)$ are narrow) it is presumed that the confidence level is comparatively high, and therefore, a relatively high energy may be allocated to the corresponding frequency components.

Rapid (and undesired) fluctuations of the estimated energy ratio $\hat{g}$ are avoided by means of temporally smoothing the estimated energy ratio $\hat{g}$ into a temporally smoothed energy ratio estimate $\hat{g}_{smooth}$. This can be accomplished by using a combination of a current estimation and, for instance, two previous estimations according to the expression:

$$\hat{g}_{smooth} = 0{,}5\hat{g}_n + 0{,}3\hat{g}_{n-1} + 0{,}2\hat{g}_{n-2}$$

where n represents a current segment number, n-1 a previous segment number and n-2 a still earlier segment number.

A high-band shape estimator 104b is included in the wide-band envelope estimator 104 in order to create a combination of the high-band shape and energy-ratio, which is probable for typical acoustic signals, such as speech signals. An estimated high-band envelope $\hat{y}$ is produced by conditioning the estimated energy ratio $\hat{g}$, the narrow-band shape and the degree of voicing r in narrow-band acoustic segment s.

A GMM with diagonal covariance matrices gives an MMSE estimate of the high-band shape $\hat{y}_{MMSE}$ according to the expression:

$$\begin{aligned}\hat{y}_{MMSE} &= E[Y \mid X = x, R = r, G = \hat{g}] \\ &= \sum_{m=1}^{M} \frac{\alpha_m f_{XRG}(x, r, \hat{g} \mid \theta_m)\mu_{y_m}}{\sum_{n=1}^{N} \alpha_n f_{XRG}(x, r, \hat{g} \mid \theta_n)}\end{aligned}$$

The excitation extension unit 105 receives the narrow-band acoustic signal a_{NB} and, on basis thereof, produces an extended excitation signal E_{WB}. As mentioned earlier, Figure 3 shows an example spectrum A_{NB} of an acoustic source signal a_{source} after having been passed through a narrow-band channel that has a bandwidth W_{NB}.

Basically, the extended excitation signal E_{WB} is generated by means of spectral folding of a corresponding excitation signal E_{NB} for the narrow-band acoustic signal a_{NB} around a particular frequency. In order to ensure a sufficient energy in a frequency region closest above the upper band limit f_{Nu} of the narrow-band acoustic signal a_{NB}, a part of the narrow-band excitation spectrum E_{NB} between a first frequency f_1 and a second frequency f_2 (where $f_1<f_2<f_{Nu}$) is cut out, e.g f_1 = 2kHz and f_2 = 3kHz, and repeatedly up-folded around first f_2, then $2f_2-f_1$, $3f_2-2f_1$ etc as many times as is necessary to cover at least the entire band up to the upper-most band limit f_{Wu}. Hence, a wide-band excitation spectrum E_{WB} is obtained. According to a preferred embodiment of the invention, the obtained excitation spectrum E_{WB} is produced such that it smoothly evolves to a white noise spectrum. This namely avoids an overly periodic excitation at the higher frequencies of the wide-band excitation spectrum E_{WB}. For instance, the transition between the up-folded narrow-band excitation spectrum E_{NB} may be set such that at the frequency $f \doteq 6$ kHz the noise spectrum dominates totally over the periodic spectrum. It is preferable, however not necessary, to allocate an amplitude of the wide-band excitation spectrum E_{WB} being equal to the mean value of the amplitude of the narrow-band excitation spectrum E_{NB}. According to an embodiment of the invention, the transition frequency depends on the confidence level for the higher frequency components, such that a comparatively high degree of certainty for these components result in a relatively high transition frequency, and conversely, a comparatively low degree of certainty for these components result in a relatively low transition frequency.

The high band shape estimator 106a in the wide-band filter 106 receives the estimated high-band envelope $\hat{y}$ from the high band shape estimator 104b and receives the wide-band excitation spectrum E_{WB} from the excitation extension unit 105. On basis of the received signals $\hat{y}$ and E_{WB}, the high band shape estimator 106a produces a high-band envelope spectrum S_Y that is shaped with the estimated high-band envelope $\hat{y}$. This frequency shaping of the excitation is performed in the frequency domain by (i) computing the wide-band excitation spectrum E_{WB} (ii) multiplying the high-band part thereof with a spectrum S_Y of the estimated high-band envelope $\hat{y}$. The high-band envelope spectrum S_Y is computed as:

$$S_Y = 10^{\frac{FFT(\hat{y}_{MMSE})}{20}}$$

A multiplier 106b receives the high-band envelope spectrum S_Y from the high band shape estimator 106a and receives the temporally smoothed energy ratio estimate $\hat{g}_{smooth}$ from the energy ratio estimator 104a. On basis of the received signals S_Y and $\hat{g}_{smooth}$ the multiplier 106b generates a high-band energy y_0. The high-band energy y_0 is determined by computing a first LFCC using only a high-band part of the spectrum between f_{Nu} and f_{Wu} (where e.g. f_{Nu} = 3,3 kHz and f_{Wu} = 8,0 kHz). The high-band energy y_0 is adjusted such that it satisfies the equation:

$$y_0 = \hat{g}_{smooth} + c_0$$

where c_0 is the energy of the current narrow-band segment (computed by the feature extraction unit 101) and $\hat{g}_{smooth}$ is the energy ratio estimate (produced by the energy ratio estimator 104a).

The high-pass filter 107 receives the high-band energy signal y_0 from the high-band shape reconstruction unit 106 and produces in response thereto a high-pass filtered signal $HP(y_0)$. Preferably, the high-pass filter's 107 cut-off frequency is set to a value above the upper bandwidth limit f_{Nu} for the narrow-band

acoustic signal a_{NB}, e.g. 3,7 kHz. The stop-band may be set to a frequency in proximity of the upper bandwidth limit f_{Nu} for the narrow-band acoustic signal a_{NB}, e.g. 3,3 kHz, with an attenuation of –60 dB.

The up-sampler 102 receives the narrow-band acoustic signal a_{NB} and produces, on basis thereof, an up-sampled signal a_{NB-u} that has a sampling rate, which matches the bandwidth W_{WB} of the wide-band acoustic signal a_{WB} that is being delivered via the signal decoder's output. Provided that the up-sampling involves a doubling of the sampling frequency, the up-sampling can be accomplished simply by means of inserting a zero valued sample between each original sample in the narrow-band acoustic signal a_{NB}. Of course, any other (non-2) up-sampling factor is likewise conceivable. In that case, however, the up-sampling scheme becomes slightly more complicated. Due to the aliasing effect of the up-sampling, the resulting up-sampled signal a_{NB-u} must also be low-pass filtered. This is performed in the following low-pass filter 103, which delivers a low-pass filtered signal $LP(a_{NB-u})$ on its output. According to a preferred embodiment of the invention, the low-pass filter 103 has an approximate attenuation of –40 dB of the high-band W_{HB}.

Finally, the adder 108 receives the low-pass filtered signal $LP(a_{NB-u})$, receives the high-pass filtered signal $HP(y_0)$ and adds the received signals together and thus forms the wide-band acoustic signal a_{WB}, which is delivered on the signal decoder's output.

In order to sum up, a general method of producing a wide-band acoustic signal on basis of a narrow-band acoustic signal will now be described with reference to a flow diagram in figure 9.

A first step 901 receives a segment of the incoming narrow-band acoustic signal. A following step 902, extracts at least one essential attribute from the narrow-band acoustic signal, which is to form a basis for estimated parameter values of a

corresponding wide-band acoustic signal. The wide-band acoustic signal includes wide-band frequency components outside the spectrum of the narrow-band acoustic signal (i.e. either above, below or both).

A step 903 then determines a confidence level for each wide-band frequency component. Either a specific confidence level is assigned to (or associated with) each wide-band frequency component individually, or a particular confidence level refers collectively to two or more wide-band frequency components. Subsequently, a step 904 investigates whether a confidence level has been allocated to all wide-band frequency components, and if this is the case, the procedure is forwarded to a step 909. Otherwise, a following step 905 selects at least one new wide-band frequency component and allocates thereto a relevant confidence level. Then, a step 906 examines if the confidence level in question satisfies a condition Γ_h for a comparatively high degree of certainty (according to any of the above-described methods). If the condition Γ_h is fulfilled, the procedure continues to a step 908 in which a relatively high parameter value is allowed to be allocated to the wide-band frequency component(s) and where after the procedure is looped back to the step 904. Otherwise, the procedure continues to a step 907 in which a relatively low parameter value is allowed to be allocated to the wide-band frequency component(s) and where after the procedure is looped back to the step 904.

The step 909 finally produces a segment of the wide-band acoustic signal, which corresponds to the segment of the narrow received that was received in the step 901.

Naturally, all of the process steps, as well as any sub-sequence of steps, described with reference to the figure 9 above may be carried out by means of a computer program being directly loadable into the internal memory of a computer, which includes appropriate software for performing the necessary steps when

the program is run on a computer. The computer program can likewise be recorded onto arbitrary kind of computer readable medium.

The term "comprises/comprising" when used in this specification is taken to specify the presence of stated features, integers, steps or components. However, the term does not preclude the presence or addition of one or more additional features, integers, steps or components or groups thereof.

The invention is not restricted to the described embodiments in the figures, but may be varied freely within the scope of the claims.

Claims

1. A method of producing a wide-band acoustic signal (a_{WB}) based on a narrow-band acoustic signal (a_{NB}), the spectrum (A_{WB}) of the wide-band acoustic signal (a_{WB}) having a larger bandwidth than the spectrum (A_{NB}) of the narrow-band acoustic signal (a_{NB}), the method involving

extraction of at least one essential attribute ($z_{NB}(r, c)$, E_{NB}) from the narrow-band acoustic signal (a_{NB}), and

estimation of a parameter describing aspects of wide-band frequency components outside the spectrum (A_{NB}) of the narrow-band acoustic signal (a_{NB}) based on at least one essential attribute ($z_{NB}(r, c)$, E_{NB}), **characterised by** allocating a parameter value to a particular wide-band frequency component based on a corresponding confidence level.

2. A method according to claim 1, **characterised by** allocating the parameter value such that

a relatively high parameter value is allowed to be allocated to the frequency component if the confidence level indicates a comparatively high degree of certainty, and

a relatively low parameter value is allowed to be allocated to the frequency component if the confidence level indicates a comparatively low degree of certainty.

3. A method according to any one of the claims 1 or 2, **characterised by** the parameter value representing a signal energy.

4. A method according to any one of the claims 1-3, **characterised by** the spectrum (A_{WB}) of the wide-band acoustic signal (a_{WB}) comprising

a low-band (W_{LB}) including wide-band frequency components below a lower bandwidth limit (f_{Nl}) of the spectrum (A_{NB}) of the narrow-band acoustic signal (a_{NB}), and

a high-band (W_{HB}) including wide-band frequency components above an upper bandwidth limit (f_{Nu}) of the spectrum (A_{NB}) of the narrow-band acoustic signal (a_{NB}),

the method involving allocating a confidence level that represents a high degree certainty to all frequency components in the low-band (W_{LB}).

5. A method according to any one of the claims 1-4, **characterised by**

receiving the narrow-band acoustic signal (a_{NB}) and on basis thereof producing an up-sampled signal ($a_{NB\text{-}u}$) having a sampling rate that matches the bandwidth (W_{WB}) of the wide-band acoustic signal (a_{WB}), and

low-pass filtering the up-sampled signal ($a_{NB\text{-}u}$) into a low-pass filtered signal ($LP(a_{NB\text{-}u})$).

6. A method according to claim 5, **characterised by** the producing of the up-sampled signal ($a_{NB\text{-}u}$) involving insertion of zero valued samples between samples of the narrow-band acoustic signal (a_{NB}).

7. A method according to any one of the claims 4-6, **characterised by** involving estimating a wide-band envelope ($\hat{s}_e$) on basis of at least one essential attribute ($z_{NB}(r, c)$).

8. A method according to claim 7, **characterised by** involving extending an excitation (E_{NB}) of the narrow-band acoustic signal (a_{NB}), the extension involving at least one spectral folding of a fraction ($f_1 - f_2$) of an excitation spectrum (E_{NB}) of the narrow-band acoustic signal (a_{NB}).

9. A method according to claim 8, **characterised by** involving wide-band filtering of the extended excitation spectrum (E_{WB}) into a wide-band energy signal (y_0), the wide-band filtering being based on the wide-band envelope estimation ($\hat{s}_e$).

10. A method according to claim 9, **characterised by** involving high-pass filtering of the wide-band energy signal (y_0) into a high-pass filtered signal ($HP(y_0)$).

11. A method according to claim 10, **characterised by** involving receiving the high-pass filtered signal ($HP(y_0)$), receiving the low-pass filtered signal ($LP(a_{NB-u})$) and producing the wide-band acoustic signal (a_{WB}) as the sum of the received signals.

12. A method according to any one of the proceeding claims, **characterised by** the at least one essential attribute ($z_{NB}(r, c)$) represents a degree of voicing and a spectral envelope (c).

13. A method according to claim 12, **characterised by** the degree of voicing being determined by a normalised auto-correlation function.

14. A method according to any one of the claims 12 or 13, **characterised by** the spectral envelope (c) being represented by means of linear frequency cepstral coefficients.

15. A method according to any one of the claims 12 or 13, **characterised by** the spectral envelope being represented by means of line spectral frequencies.

16. A method according to any one of the claims 12 or 13, **characterised by** the spectral envelope being represented by means of Mel frequency cepstral coefficients.

17. A method according to any one of the claims 12 or 13, **characterised by** the spectral envelope being represented by means of linear prediction coefficients.

18. A method according to any one of the claims 7-17, **characterised by** the estimation of the high-band (W_{HB}) fraction of the wide-band envelope ($\hat{s}_e$) involving Gaussian mixture modelling.

19. A method according to claim 18, **characterised by** the Gaussian mixture modelling involving

Bayes classification of at least one narrow-band feature vector into a mixture component of a Gaussian mixture model, and

computation of a value that indicates the probability of that the classification is correct.

20. A method according to claim 18, **characterised by** the Gaussian mixture model representing a joint distribution of feature vectors and underlying parameters.

21. A method according to any one of the claims 7-17, **characterised by** the estimation of the high-band (W_{HB}) fraction of the wide-band envelope ($\hat{s}_e$) involving hidden Markov modelling.

22. A computer program directly loadable into the internal memory of a computer, comprising software for performing the

steps of any of the claims 1–21 when said program is run on the computer.

23. A computer readable medium, having a program recorded thereon, where the program is to make a computer perform the steps of any of the claims 1–21.

24. A signal decoder for producing a wide-band acoustic signal (a_{WB}) from a narrow-band acoustic signal (a_{NB}), the spectrum (A_{WB}) of the wide-band acoustic signal (a_{WB}) having a larger bandwidth than the spectrum (A_{NB}) of the narrow-band acoustic signal (a_{NB}), the signal decoder comprising:

a feature extraction unit (101) receiving the narrow-band acoustic signal (a_{NB}) and on basis thereof producing at least one essential attribute ($z_{NB}(r, c)$, E_{NB}) of the narrow-band acoustic signal (a_{WB}), and

at least one band extension unit (102 - 108) receiving the narrow-band acoustic signal (a_{NB}), receiving the at least one essential attribute ($z_{NB}(r, c)$, E_{NB}) and on basis of the received signals producing the wide-band acoustic signal (a_{WB}),

characterised in that

the signal decoder is arranged to allocate a parameter with respect to a particular wide-band frequency component based a corresponding confidence level.

25. A signal decoder according to claim 24, **characterised in that** the signal decoder is arranged to allocate the parameter such that

a relatively high parameter value is allowed to be allocated to the frequency component if the confidence level indicates a comparatively high degree certainty, and

a relatively low parameter value is allowed to be allocated to the frequency component if the confidence level indicates a comparatively low degree certainty.

26. A signal decoder according to claim 24 or 25, **characterised in that** the parameter value represents a signal energy.

27. A signal decoder according to any one of the claims 24-26, **characterised in that** it comprises

an up-sampler (102) receiving the narrow-band acoustic signal (a_{NB}) and on basis thereof producing an up-sampled signal (a_{NB-u}) that has a sampling rate, which matches the bandwidth (W_{WB}) of the wide-band acoustic signal (a_{WB}), and

a low-pass filter (103) receiving the up-sampled signal (a_{NB-u}) and in response thereto producing a low-pass filtered acoustic signal ($LP(a_{NB-u})$).

28. A signal decoder according to any one of the claims 24-27, **characterised in that** it comprises a wide-band envelope estimator (104) receiving the at least one essential attribute ($z_{NB}(r, c)$) and on basis thereof producing an estimated wide-band envelope ($\hat{s}_e$).

29. A signal decoder according to claim 28, **characterised in that** the wide-band envelope estimator (104) comprises an energy ratio estimator (104a) receiving the at least one essential attribute ($z_{NB}(r, c)$) and in response thereto producing an estimated energy ratio ($\hat{g}$).

30. A signal decoder according to claim 29, **characterised in that** the wide-band envelope estimator (104) comprises a high-band shape estimator (104b) receiving the at least one essential attribute ($z_{NB}(r, c)$), receiving the estimated energy ratio ($\hat{g}$) and on basis of the received signals producing an estimated high-band envelope ($\hat{y}$).

31. A signal decoder according to any one of the claims 28-30, **characterised in that** it comprises an excitation extension unit (105) receiving the narrow-band acoustic signal (a_{NB}) and in response thereto producing an extended excitation spectrum (E_{WB}), the extended excitation spectrum (E_{WB}) comprising frequency components outside the spectrum (A_{NB}) of the narrow-band acoustic signal (a_{NB}).

32. A signal decoder according to claim 31, **characterised in that** it comprises a wide-band filter (106) receiving the extended excitation spectrum (E_{WB}), receiving the wide-band envelope estimation ($\hat{s}_e$) and on basis of the received signals producing a wide-band energy signal (y_0).

33. A signal decoder according to claim 32, **characterised in that** the wide-band filter (106) comprises a high-band shape-reconstruction unit (106a) receiving the extended excitation spectrum (E_{WB}), receiving the estimated high-band envelope ($\hat{y}$) and on basis of the received signals producing a high-band envelope spectrum (S_Y).

34. A signal decoder according to claim 33, **characterised in that**

the energy ratio estimator (104a) comprises means for producing a temporally smoothed energy ratio estimate ($\hat{g}_{smooth}$) on basis of the at least one essential attribute ($z_{NB}(r, c)$), and

the wide-band filter (106) comprises a multiplier (106b) receiving the high-band envelope spectrum (S_Y), receiving the temporally smoothed energy ratio estimate ($\hat{g}_{smooth}$) and on basis of the received signals producing the wide-band energy signal (y_0).

35. A signal decoder according to any one of the claims 31-34, **characterised in that** it comprises a high-pass filter (107) receiving the wide-band energy signal (y_0) and in response thereto producing a high-pass filtered signal ($HP(y_0)$).

36. A signal decoder to claim 35, **characterised in that** it comprises an adder (108) receiving the high-pass filtered signal ($HP(y_0)$), receiving the low-pass filtered signal ($LP(a_{NB-u})$) and producing the wide-band acoustic signal (a_{WB}) as a sum of the received signals.

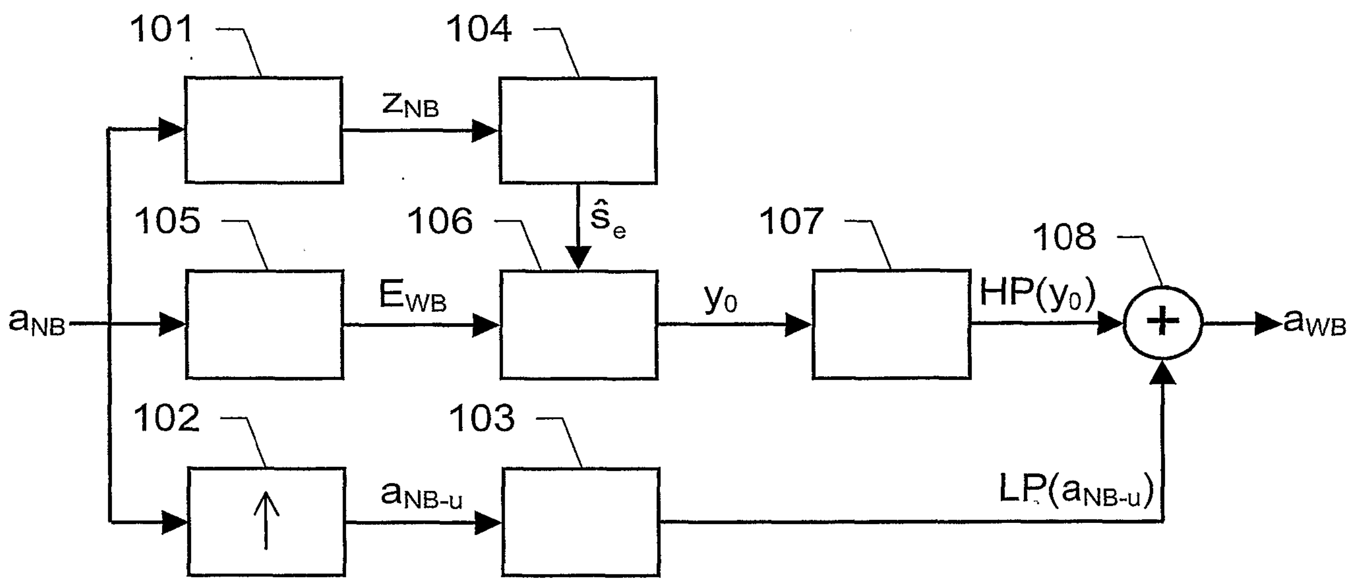

101
104
z_{NB}
105
106
$\hat{s}_e$
107
108
a_{NB}
E_{WB}
y_0
$HP(y_0)$
a_{WB}
102
103
a_{NB-u}
$LP(a_{NB-u})$

Fig. 1

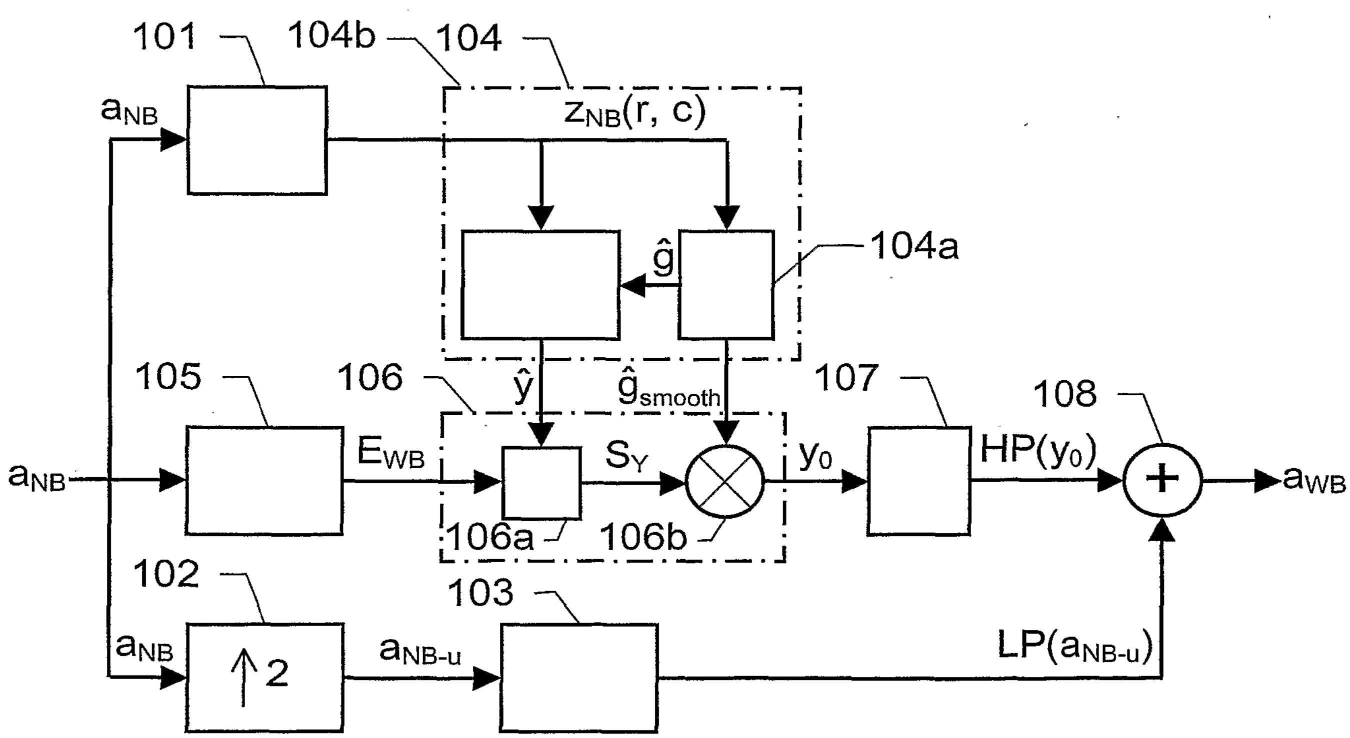

101
104b
104
a_{NB}
$z_{NB}(r, c)$
$\hat{g}$
104a
105
106
$\hat{y}$
$\hat{g}_{smooth}$
107
108
a_{NB}
E_{WB}
S_Y
y_0
$HP(y_0)$
a_{WB}
106a
106b
102
103
a_{NB}
2
a_{NB-u}
$LP(a_{NB-u})$

Fig. 5

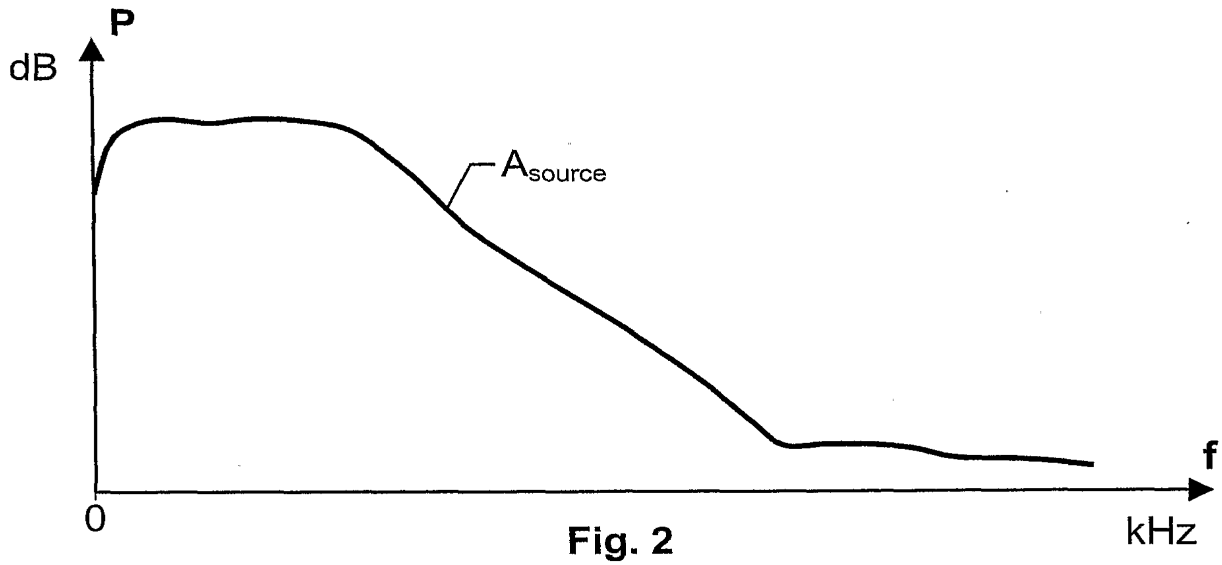


Fig. 2

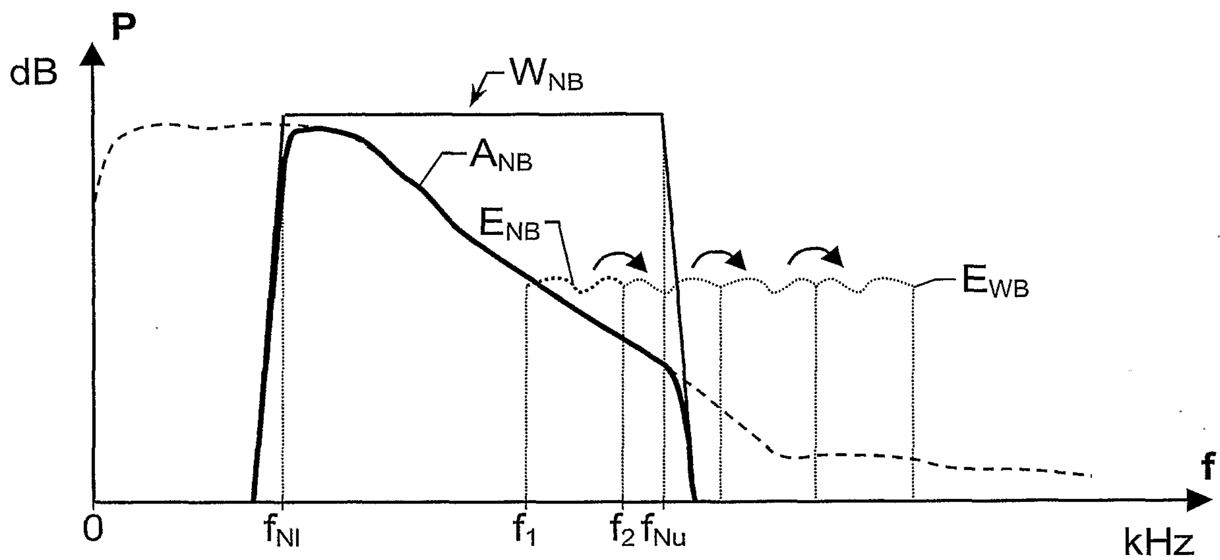


Fig. 3

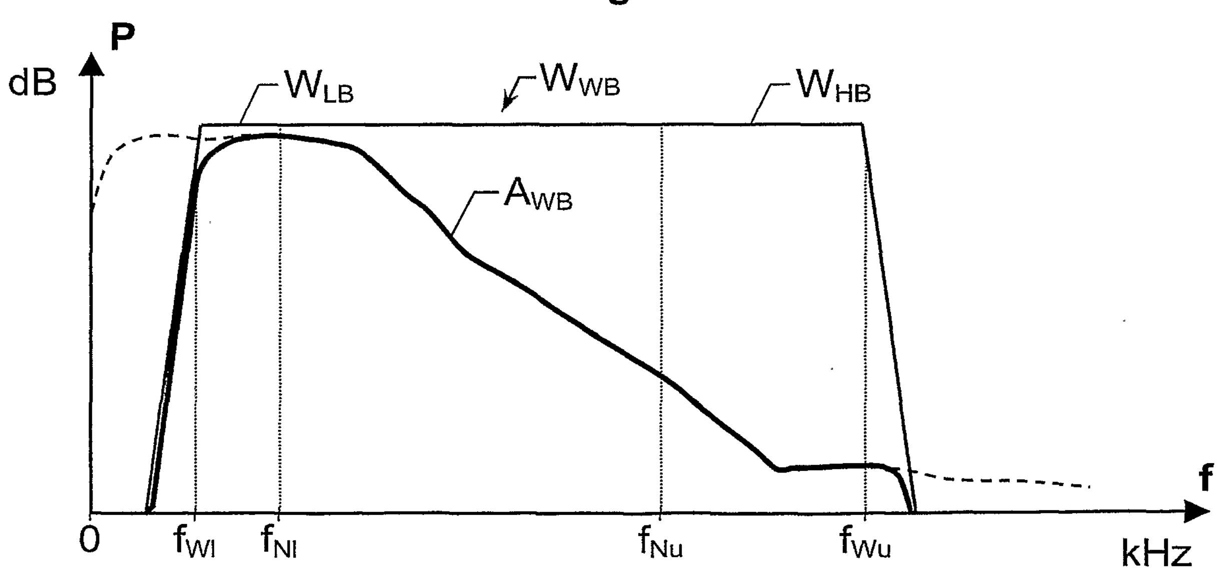


Fig. 4

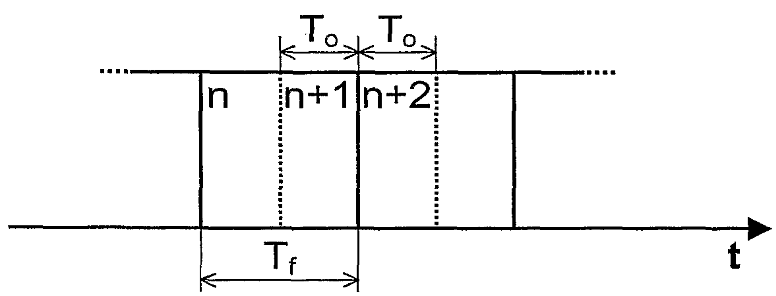


Fig. 6

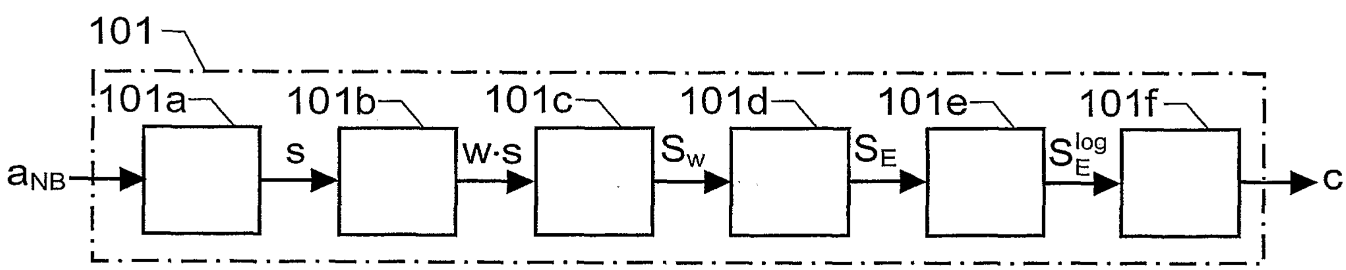


Fig. 7

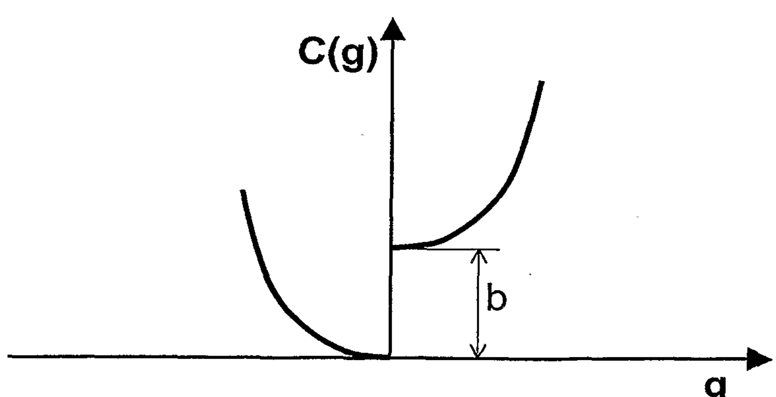


Fig. 8

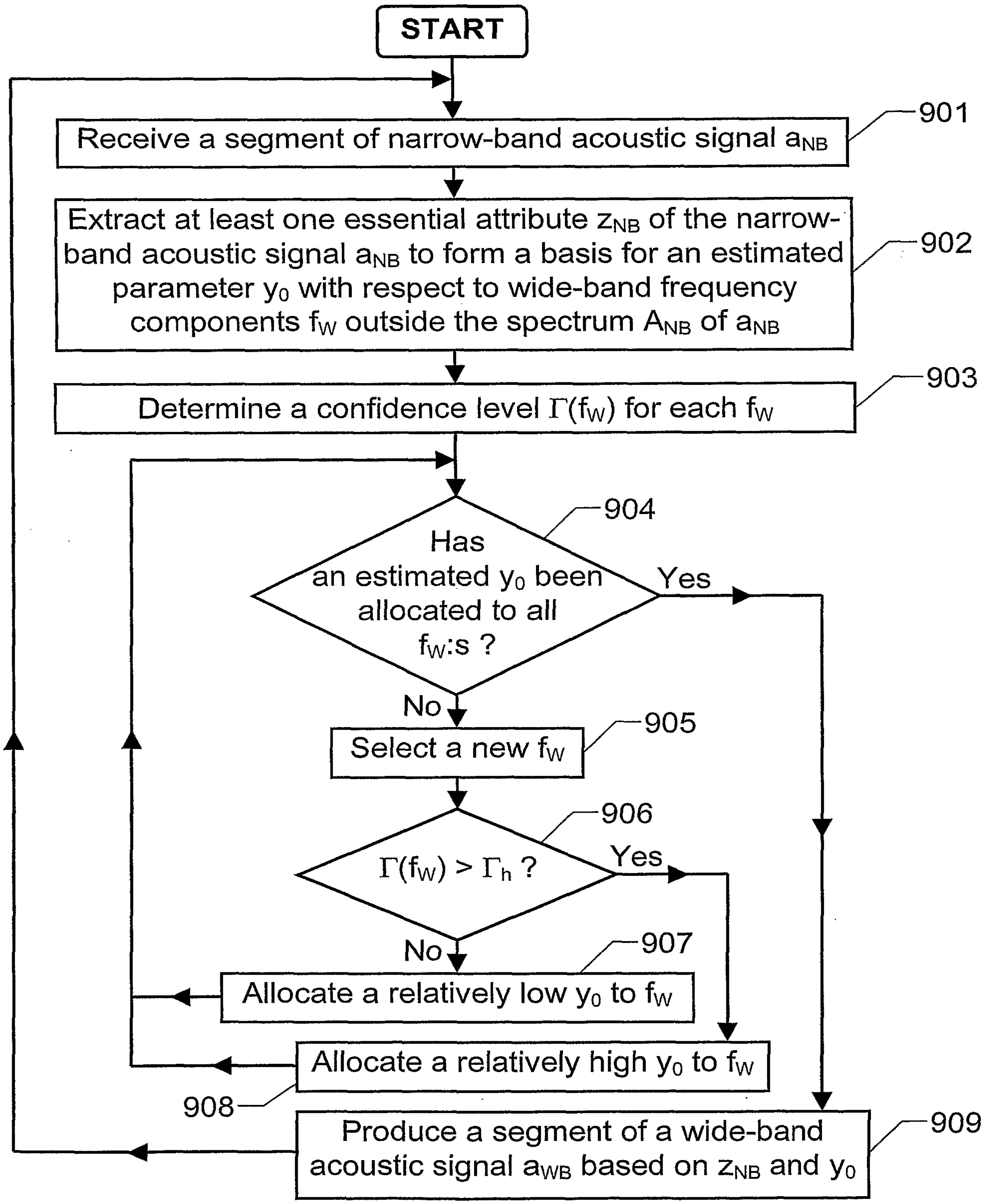

START
Receive a segment of narrow-band acoustic signal a_{NB}
901
Extract at least one essential attribute z_{NB} of the narrow-band acoustic signal a_{NB} to form a basis for an estimated parameter y_0 with respect to wide-band frequency components f_W outside the spectrum A_{NB} of a_{NB}
902
Determine a confidence level $\Gamma(f_W)$ for each f_W
903
Has an estimated y_0 been allocated to all f_W:s ?
904
Yes
No
Select a new f_W
905
$\Gamma(f_W) > \Gamma_h$?
906
Yes
No
Allocate a relatively low y_0 to f_W
907
Allocate a relatively high y_0 to f_W
908
Produce a segment of a wide-band acoustic signal a_{WB} based on z_{NB} and y_0
909

Fig. 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 02/00485

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: G10L 19/02 // G10L 21/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-INTERNAL, WPI DATA

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5956686 A (TAKASHIMA, M. ET AL.), 21 Sept 1999 (21.09.99), column 6, line 31 - column 7, line 13, abstract, the whole document	1-33,35-36
A		34
	--	
A	WO 0103124 A1 (TELEFONAKTIEBOLAGET LM ERICSSON), 11 January 2001 (11.01.01)	1-36
	--	
A	US 5978759 A (TSUSHIMA, M. ET AL.), 2 November 1999 (02.11.99)	1-36
	--	
A	EP 1089258 A2 (SONY CORP), 4 April 2001 (04.04.01)	1-36
	--	

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 June 2002	18 -06- 2002

Name and mailing address of the ISA/	Authorized officer
Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Henrik Andersson/LR Telephone No. +46 8 782 25 00

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 1008984 A2 (SONY CORP), 14 June 2000 (14.06.00) --	1-36
A	US 5455888 A (IYENGAR, V. ET AL.), 3 October 1995 (03.10.95), column 4, line 50 - line 57, claims 1, 2,10,11, abstract -- --------	1,2,4,7,9, 12,18,24,25, 28

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/SE 02/00485

Patent document cited in search report			Publication date	Patent family member(s)			Publication date
US	5956686	A	21/09/99	CN	1121620	A	01/05/96
				JP	3237089	B	10/12/01
				JP	8044392	A	16/02/96
WO	0103124	A1	11/01/01	AU	5818000	A	22/01/01
				GB	2351889	A	10/01/01
				GB	9915831	D	00/00/00
US	5978759	A	02/11/99	DE	69619284	D	00/00/00
				EP	0732687	A,B	18/09/96
				JP	3189614	B	16/07/01
				JP	8248997	A	27/09/96
				JP	2798003	B	17/09/98
				JP	8305396	A	22/11/96
				JP	2956548	B	04/10/99
				JP	9101798	A	15/04/97
EP	1089258	A2	04/04/01	CN	1297222	A	30/05/01
				JP	2001100773	A	13/04/01
EP	1008984	A2	14/06/00	CN	1261713	A	02/08/00
				JP	2000181494	A	30/06/00
				JP	2000181495	A	30/06/00
				JP	2000181496	A	30/06/00
				JP	2000181497	A	30/06/00
				JP	2000206997	A	28/07/00
US	5455888	A	03/10/95	NONE			